national accelerator laboratory

FN-243
2042

Second Order Contributions to Beam Dimensions

David C. Carey

May, 1972

I. Introduction

The phase space region occupied by an aggregate of charged particles in a beam line is often represented by a higher dimensional ellipsoid. Given no further information, one might interpret such an ellipsoid as an envelope inside of which particles are distributed uniformly, or as giving the scale dimensions of a gaussian distribution. The latter case has the advantage that is easily adapted to include higher order effects of the beam line. In either case the parameters of the ellipsoid are simply related to the first and second moments and therefore the width of the distribution in any coordinate. In first order an ellipsoid at any point in a beam line is transformed into another ellipsoid at any other location in a beam line. In second and higher orders a transformation from one location in a beam line to another will cause the ellipsoid to become distorted. One can still, however, calculate the first and second moments of the distribution, and thereby obtain a measure of its dimensions in any coordinate.

Below we elaborate on the methods for calculating the ellipsoid parameters at any point in the beam line. Much of the first order theory can be found in the work of Brown and Howry.[1] It is included here for completeness.

Operated by Universities Research Association Inc. Under Contract with the United States Atomic Energy Commission

II. The Ellipsoid Formalism

The position and motion of a particle in a beam line may be represented via a six-dimensional vector.

$$x = \begin{pmatrix} x \\ \theta \\ y \\ \phi \\ \ell \\ \delta \end{pmatrix} \tag{1}$$

The coordinates x and y represent respectively the horizontal and vertical displacements at the position of the particle, θ and ϕ, the angles with the axis of the beam line in the same planes. The quantity ℓ represents the longitudinal position of the particle relative to a particle traveling on the magnetic axis of the system with the central momentum designed for the system. The remaining quantity $\delta = \frac{\Delta p}{p}$ gives the fractional deviation of the momentum of the particle from the central design momentum of the system.

An ellipsoidal hypersurface in this six-dimensional space may be represented by the equation:

$$x^T \sigma^{-1} x = 1 \tag{2}$$

where σ^{-1} is a symmetric positive definite matrix. We represent this matrix as an inverse for reasons which will become apparent later. At this stage the center of the ellipsoid is assumed to lie at the origin of the coordinate system. The ellipsoid may be taken to be the envelope of a uniform distribution, or the scale in a gaussian distribution, giving a particle density:

$$\rho = C \exp(-\tfrac{1}{2} x^T \sigma^{-1} x) \tag{3}$$

For any real symmetric matrix there exists a coordinate system in which that matrix is diagonal and an orthogonal transformation to that coordinate system.[2] Let us represent the orthogonal transformation by the matrix O, so that:

$$x_i = \sum_j O_{ij} \tilde{x}_j \tag{4}$$

where $\tilde{x}_j$ are the coordinates in the frame where the transform of σ^{-1} and therefore that of σ are diagonal. Calling the matrix σ transformed to the new frame $\tilde{\sigma}$ we now have:

$$\sigma_{ij} = \sum_{k\ell} O_{ik} \tilde{\sigma}_{k\ell} O_{j\ell} \tag{5}$$

and equation (1) becomes

$$\tilde{x}^T \tilde{\sigma}^{-1} \tilde{x} = 1 \tag{1a}$$

Specializing to the gaussian distribution, it is now an easy matter to calculate the second moments in the new frame since the coordinates are decoupled. We arrive at:

$$\overline{\tilde{x}_i \tilde{x}_j} = \tilde{\sigma}_{ij} = \delta_{ij} \tilde{\sigma}_{jj} \tag{6}$$

The second moments in the old frame are now:

$$\begin{aligned} \overline{x_i x_j} &= \overline{\sum_{k\ell} O_{ik} O_{j\ell} \tilde{x}_k \tilde{x}_\ell} = \sum_{k\ell} O_{ik} O_{j\ell} \overline{\tilde{x}_k \tilde{x}_\ell} \\ &= \sum_{k\ell} O_{ik} O_{j\ell} \tilde{\sigma}_{k\ell} = \sigma_{ij}. \end{aligned} \tag{7}$$

Therefore in this case the elements of the matrix σ give the second moments of the distribution in the original coordinate system. The density function, properly normalized, now becomes:

$$\rho = \frac{N_o}{\sqrt{\det(\sigma)}\ (2\pi)^3} \exp(-\tfrac{1}{2} x^T \sigma^{-1} x) \tag{8}$$

where N_o is the total number of particles. Since the matrix O is orthogonal the determinants of σ and $\tilde{\sigma}$ are equal.

The elements of the matrix σ may be put in more convenient form for interpretation. The square roots of the diagonal elements may be taken as giving the half widths x_o of the distribution in a given coordinate while the off-diagonal elements may be related to the correlations r_{ij}, so

$$x_{oi} = \sqrt{\sigma_{ii}}$$

$$r_{ij} = \sigma_{ij}/\sqrt{\sigma_{ii}\sigma_{jj}} \tag{9}$$

Since, for any positive definite symmetric matrix σ, we have:[2]

$$\sigma_{ii}\ \sigma_{jj} - \sigma_{ij}^{\ 2} > 0 \tag{10}$$

the correlations must all obey the inequality

$$|r_{ij}| < 1 \tag{11}$$

If the ellipsoid is interpreted as describing the envelope of a uniform distribution, then the x_{oi} represent the maximum extents of the beam in the given coordinates.

III. The Effect of a Beam Line

A. First Order

If we now let $x_j^{(1)}$ be the coordinates of a ray at the initial point in a beam line, and $x_i^{(2)}$ the coordinates at some later point, the two are related by the equation:

$$x_i^{(2)} = \sum_j R_{ij}\ x_j^{(1)} \tag{12}$$

If we continue to assume a distribution centered at the origin the first moments at both initial and final point will be zero. The second moments will now be given by:

$$\sigma_{ij}^{(2)} = \overline{x_i^{(2)} \; x_j^{(2)}} = \sum_{k\ell} R_{ik} \; R_{j\ell} \; \overline{x_k^{(1)} \; x_\ell^{(1)}} \tag{13}$$

$$= \sum_{k\ell} R_{ik} \; R_{j\ell} \; \sigma_{k\ell}^{(1)}$$

or more concisely

$$\sigma^{(2)} = R\sigma^{(1)} R^T \tag{14}$$

To first order an ellipsoid at the initial point will transform into an ellipsoid at the final point, so that the equation:

$$x^{(2)^T} (\sigma^{(2)})^{-1} \; x^{(2)} = 1 \tag{15}$$

will give the envelope of the particle distribution at the later point.

B. Second Order

In second order the transformation on the coordinates effected by the beam line is given by:

$$x_i^{(2)} = \sum_j R_{ij} \; x_j^{(1)} + \sum_{jk} T_{ijk} \; x_j^{(1)} \; x_k^{(1)} . \tag{16}$$

We employ here a symmetric T matrix whose off-diagonal elements are half those of the T matrix used by Brown. The first and second moments of the distribution at the final point are now given by:

$$\overline{x_i^{(2)}} = \sum_j R_{ij} \; \overline{x_j^{(1)}} + \sum_{jk} T_{ijk} \; \overline{x_j^{(1)} \; x_k^{(1)}} \tag{17}$$

$$\overline{x_i^{(2)} \; x_j^{(2)}} = \sum_{k\ell} R_{ik} \; R_{j\ell} \; \overline{x_k^{(1)} \; x_\ell^{(1)}}$$

$$+ \sum_{k\ell m} \left[R_{ik} \; T_{j\ell m} + T_{ik\ell} \; R_{jm} \right] \overline{x_k^{(1)} \; x_\ell^{(1)} \; x_m^{(1)}}$$

$$+ \sum_{k\ell mn} T_{ik\ell} \; T_{jmn} \; \overline{x_k^{(1)} \; x_\ell^{(1)} \; x_m^{(1)} \; x_n^{(1)}}$$

For a symmetric, on-axis initial distribution, the first and third moments vanish. The problem now reduces to determining the fourth moments of the initial distribution.

As an extension of previous notation we now denote the fourth moments of the distribution about the initial point by $\sigma^{(1)}_{ijk\ell}$. We consider the coordinate system in which the matrix of second moments σ_{ij} is diagonalized, denoting the moments in this frame by $\tilde{\sigma}$. Then from equation (7) we have:

$$\sigma_{ij} = \sum_{k\ell} O_{ik}\, O_{j\ell}\, \tilde{\sigma}_{k\ell} \tag{18}$$

$$= \sum_{k} O_{ik}\, O_{jk}\, \tilde{\sigma}_{kk}$$

We continue to specialize to a gaussian distribution so that the fourth moments will be directly derviable from the second moments. In the diagonal frame the coordinates separate, and the fourth moments are easily calculated. The only ones which are non-zero are $\tilde{\sigma}_{iijj}$, $\tilde{\sigma}_{ijij}$, or $\tilde{\sigma}_{ijji}$ for $i \neq j$, and $\tilde{\sigma}_{iiii}$ with:

$$\begin{aligned}
\tilde{\sigma}_{iijj} &= \tilde{\sigma}_{ii}\, \tilde{\sigma}_{jj} \\
\tilde{\sigma}_{ijij} &= \tilde{\sigma}_{ii}\, \tilde{\sigma}_{jj} \\
\tilde{\sigma}_{ijji} &= \tilde{\sigma}_{ii}\, \tilde{\sigma}_{jj} \\
\tilde{\sigma}_{iiii} &= 3\, \tilde{\sigma}_{ii}\, \tilde{\sigma}_{ii}
\end{aligned} \tag{19}$$

so that in general:

$$\begin{aligned}
\tilde{\sigma}_{ijk\ell} = {} & \delta_{ij}\, \delta_{k\ell}\, \tilde{\sigma}_{ii}\, \tilde{\sigma}_{kk} + \delta_{ik}\, \delta_{j\ell}\, \tilde{\sigma}_{ii}\, \tilde{\sigma}_{jj} \\
& + \delta_{i\ell}\, \delta_{jk}\, \tilde{\sigma}_{ii}\, \tilde{\sigma}_{jj}.
\end{aligned} \tag{20}$$

Now if under the transformation O, the fourth moments transform as:

$$\sigma_{ijk\ell} = \sum_{mnop} O_{im}\, O_{jn}\, O_{ko}\, O_{\ell p}\, \tilde{\sigma}_{mnop} \tag{21}$$

then from equation (1a) we finally arrive at:

$$\sigma_{ijk\ell} = \sigma_{ij}\, \sigma_{k\ell} + \sigma_{ik}\, \sigma_{j\ell} + \sigma_{i\ell}\, \sigma_{jk} \tag{22}$$

Substituting into equation (17) we determine that:

$$\overline{x_i^{(2)}} = \sum_{jk} T_{ijk}\, \sigma_{jk}^{(1)} \tag{23}$$

$$\overline{x_i^{(2)}\ x_j^{(2)}} = \sum_{k\ell} R_{ik}\, R_{j\ell}\, \sigma_{k\ell}^{(1)} + (\sum_{k\ell} T_{ik\ell}\, \sigma_{k\ell}^{(1)})\ (\sum_{mn} T_{jmn}\, \sigma_{mn}^{(1)})$$

$$+ 2 \sum_{\ell m} (\sum_{k} T_{ik\ell}\, \sigma_{km}^{(1)})\ (\sum_{n} T_{jmn}\, \sigma_{\ell n}^{(1)})$$

Note that, because of the symmetry properties of both T and σ that the two expressions in parentheses in the last term of the second equation represent the same array. From a practical standpoint this means that it needs to be calculated only once.

We see from equation (23) that the centroid of the distribution at the final point no longer coincides with the beam axis. Letting $\sigma^{(2)}$ represent the matrix of second moments about the new centroid we now have:

$$\sigma_{ij}^{(2)} = \overline{x_i^{(2)}\ x_j^{(2)}} - \overline{x_i^{(2)}}\ \overline{x_j^{(2)}} \tag{24}$$

$$= \sum_{k\ell} R_{ik}\, R_{j\ell}\, \sigma_{k\ell}^{(1)}$$

$$+ 2 \sum_{\ell m} (\sum_{k} T_{ik\ell}\, \sigma_{km}^{(1)})\ (\sum_{n} T_{jmn}\, \sigma_{\ell n}^{(1)})$$

IV. Off-Axis Initial Distribution

Now consider a gaussian distribution whose center does not coincide with the beam axis. Letting the coordinates of the centroid by $\overline{x_i^{(1)}}$, we have for the coordinates of a ray:

$$x_i^{(1)} = \overline{x_i^{(1)}} + \xi_i^{(1)} \tag{25}$$

We let the matrix σ represent the moments of the distribution about its centroid so that:

$$\overline{\xi_i^{(1)}\ \xi_j^{(1)}} = \sigma_{ij}^{(1)} \tag{26}$$

$$\overline{\xi_i^{(1)}\ \xi_j^{(1)}\ \xi_k^{(1)}\ \xi_\ell^{(1)}} = \sigma_{ijk\ell}^{(1)}$$

Equation (17) continues to hold for the moments of the distribution about the b eam axis, while equation (22) holds for the moments about the centroid. We must therefore express one set of moments in terms of the other.

Using equations (22), (25), and (26) and applying the first part of equation (24) to the initial distribution, the initial third and fourth moments are given in terms of the initial first and second moments as follows:

$$\begin{aligned}
\overline{x_i^{(1)}\ x_j^{(1)}\ x_k^{(1)}} &= \overline{x_i^{(1)}}\ \overline{x_j^{(1)}\ x_k^{(1)}} + \overline{x_j^{(1)}}\ \overline{x_i^{(1)}\ x_k^{(1)}} \\
&\quad + \overline{x_k^{(1)}}\ \overline{x_i^{(1)}\ x_j^{(1)}} - 2\ \overline{x_i^{(1)}}\ \overline{x_j^{(1)}}\ \overline{x_k^{(1)}} \\
\overline{x_i^{(1)}\ x_j^{(1)}\ x_k^{(1)}\ x_\ell^{(1)}} &= \overline{x_i^{(1)}\ x_j^{(1)}}\ \overline{x_k^{(1)}\ x_\ell^{(1)}} + \overline{x_i^{(1)}\ x_k^{(1)}}\ \overline{x_j^{(1)}\ x_\ell^{(1)}} \\
&\quad + \overline{x_i^{(1)}\ x_\ell^{(1)}}\ \overline{x_j^{(1)}\ x_k^{(1)}} - 2\ \overline{x_i^{(1)}}\ \overline{x_j^{(1)}}\ \overline{x_k^{(1)}}\ \overline{x_\ell^{(1)}}
\end{aligned} \tag{27}$$

Substituting into equation (17) and rearranging terms we arrive at the following expressions for the first and second moments of the distribution at the final point.

$$\overline{x_i^{(2)}} = \sum_j R_{ij} \overline{x_j^{(1)}} + \sum_{jk} T_{ijk} \overline{x_j^{(1)} x_k^{(1)}} \tag{28}$$

$$\begin{aligned} \overline{x_i^{(2)} x_j^{(2)}} = {} & \sum_{k\ell} R_{ik} R_{j\ell} \overline{x_k^{(1)} x_\ell^{(1)}} + \overline{x_i^{(2)}}\ \overline{x_j^{(2)}} - 2 X_i^{(2)} X_j^{(2)} \\ & + 2 \sum_{\ell m} \left(R_{i\ell} \overline{x_m^{(1)}} + \sum_k T_{ik\ell} \overline{x_k^{(1)} x_m^{(1)}}\right) \left(R_{jm} \overline{x_\ell^{(1)}} + \right. \\ & \left. \sum_n T_{jmn} \overline{x_\ell^{(1)} x_n^{(1)}}\right) - \left(\sum_k R_{ik} \overline{x_k^{(1)}}\right) \left(\sum_m R_{jm} \overline{x_m^{(1)}}\right) \end{aligned}$$

where

$$X_i^{(2)} = \sum_k R_{ik} \overline{x_k^{(1)}} + \sum_{k\ell} T_{ik\ell} \overline{x_k^{(1)}}\ \overline{x_\ell^{(1)}}$$

is the image of the original centroid.

We may now again use equations (9) and (24) to relate this matrix of second moments to the final beam half widths and correlations.

References

1. Karl L. Brown, Sam K. Howry, SLAC Report No. 91 (1970).
2. F. R. Gantmacher, The Theory of Matrices, Chelsea Publishing Co., New York (1959).